OsteoArthritis and Cartilage (2004) 12, 912–916

doi:10.1016/j.joca.2004.08.006

Osteoarthritis and Cartilage

International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Responsiveness of the electronic touch screen WOMAC 3.1 OA Index in a short term clinical trial with rofecoxib[1]

R. Theiler†*, H. A. Bischoff-Ferrari‡, M. Good§ and N. Bellamy‖

† *Rheumatology Clinic and Institute for Physical Medicine and Rehabilitation, Stadtspital Triemli, Zurich, Switzerland*

‡ *Division of Rheumatology, Immunology and Allergy, Robert B. Brigham Arthritis and Musculoskeletal Clinical Research Center, Boston, MA, USA*

§ *Target BioScience Ltd., Rüschlikon-Zurich, Switzerland*

‖ *Centre of National Research on Disability and Rehabilitation Medicine (CONROD), Mayne Medical School, Floor 3, Herston RD, Brisbane, Australia*

Summary

Background: The Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index is a self-administered validated questionnaire for patients with osteoarthritis (OA) of the hip or knee. The electronic touch screen version of the WOMAC (e-WOMAC) has been previously shown to be highly correlated with the original paper format. However, whether the e-WOMAC would be suitable for monitoring the effects of drug treatment is unknown.

Aim: To validate the longitudinal use of the e-WOMAC questionnaire and its ability to detect changes in WOMAC-scores induced by drug treatment in outpatient care.

Methods: Fifty-three outpatients, men and women (mean age: 64 years; SD $\pm$ 9.5), with symptomatic osteoarthritis of hip or knee were included in an open label study with rofecoxib. At three visits over 3 weeks, responsiveness of the WOMAC 3.1 regarding the three subscales, pain, stiffness and function, were compared for the original paper format and the computer touch screen format (QUALITOUCH®) using a Likert scale. WOMAC scores were transformed to the 0–100 scale. ANOVA for repeated measures was used for analysis and effect sizes by subscale were compared for both formats.

Results: Responsiveness for all three subscales was similar between formats. In both formats, pain and stiffness were significantly reduced with rofecoxib as early as 7 days, while functional ability was significantly increased ($P < 0.01$ for all aggregate subscale scores) with continuing improvement until the end of study. The effect sizes by subscale between Day 1 and 21 were not statistically different between the paper and the electronic version of the questionnaire and showed similar clinically meaningful improvements in WOMAC scores over 3 weeks.

Conclusion: In this longitudinal intervention study, the e-WOMAC OA Index 3.1 showed similar responsiveness in detecting clinically meaningful changes than the original paper format.

Key words: Patient self-assessment, Electronic WOMAC 3.1, Electronic data capturing (EDC), QUALITOUCH method, Rofecoxib.

Introduction

The Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index was developed for standardized assessment of osteoarthritis (OA) symptoms in hip and/or knee joints[1]. The WOMAC OA Index covers the domains of pain, stiffness and function in 5, 2 and 17 questions, respectively. This index has been extensively validated for measuring changes after different interventions in patients with OA[2] and is the most widely recommended disease-specific questionnaire for core set assessment in clinical trials for knee and hip OA established by the Osteoarthritis Society Task Force, as proposed at the third conference on outcome measures in rheumatoid arthritis clinical trials (OMERACT III)[3].

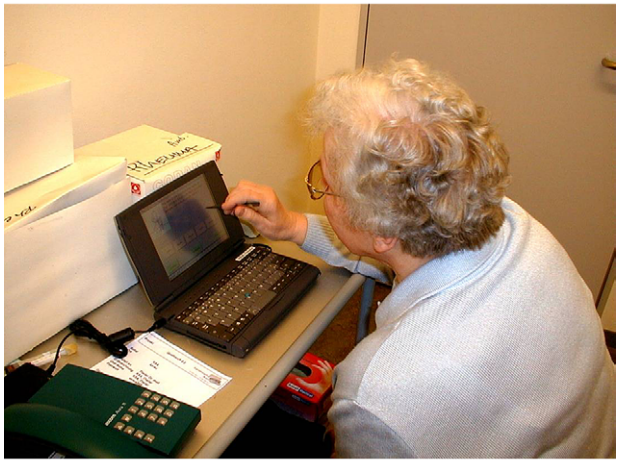
In daily clinical practice the WOMAC questionnaire is a suitable tool for optimizing patient monitoring as the data are directly provided by the patient and are very reproducible. However, the paper format does not allow for an immediate display of results. The e-WOMAC was designed to improve patient monitoring by its simple design and provides the opportunity to discuss results with the patients or the team that takes care of the patients in a timely fashion, as results are available immediately and can be shared electronically[4]. Another advantage of the e-WOMAC may be its presentation format, where each question is displayed as text and a situational cartoon, and are verbalized over the loudspeaker (QUALITOUCH®

[1]This study was supported by a grant of MSD-Chibret AG, Glattbrugg, Switzerland.

*Author correspondence and reprint requests to: PD Dr. Med. Robert Theiler, Head of Rheumatological Clinic and Institute for Physical Medicine and Rehabilitation, Stadtspital Triemli, CH-8063 Zürich, Switzerland. Tel: 41-(0)1-466-23-02.; E-mail: robert.theiler@triemli.stzh.ch

Received 23 November 2003; revision accepted 11 August 2004.

a



b

Fig. 1. (a) QUALITOUCH® multimedia 3-D interactive interface—questions answered by touching one of the squares of the Likert scale on the computer screen. (b) Screen display of the electronic touch screen featuring question 6 of the WOMAC 3.1 OA Index and the Likert scale. Translation: Steifigkeit = stiffness. Wie stark ist Ihre Gelenksteifigkeit nach dem ersten Aufwachen am Morgen? = how important is the stiffness of your joints after the first awakening in the morning. Keine = none. Leichte = mild. Mässige = moderate. Starke = severe. Extreme = extreme. Frage wiederholen = repeat question. Befragung abbrechen = stop interview. Hilfe = help. Vorherige Frage = previous question. Nächste Frage = next question.

method). This may be appreciated especially by the older patients. The electronic formats of the WOMAC-Index 3.0 have been previously validated[5–7].

Of special clinical and scientific interest is the ability of a questionnaire to detect and monitor improvement or worsening of the clinical situation based on an intervention. We therefore compared the responsiveness of the original paper format and the computer touch screen format to rofecoxib treatment in patients with symptomatic hip or knee OA over a course of 3 weeks. The aim of this study was to validate the longitudinal responsiveness of the e-WOMAC questionnaire.

Patients and methods

Three sites that participated in the previously published SVIS-Study, with 136 recruited patients and 22 sites in total[5], participated in this ancillary e-WOMAC protocol and 53 eligible consecutive outpatients were recruited and included in the evaluation. The SVIS-study was a prospective open label 3 weeks multicenter study to document the effect of rofecoxib in patients with painful radiographically proven primary OA of the knee or the hip according to ACR criteria who were dissatisfied with their prior NSAID treatment (because of either non-responsiveness to or adverse events from previous NSAID-therapy, including celecoxib). At inclusion the patients stopped their previous NSAID therapy and started therapy with rofecoxib 25 mg once daily on the following day (t_0) for 3 weeks, after which the final visit took place (t_2), with an interim visit on day 7 (t_1)[5].

Because the core study was a clinical trial with drug intervention, conducted according to GCP guidelines and the e-WOMAC is not yet a validated format of the questionnaire acceptable to regulatory authorities, we had to renounce to block randomize the patients for the paper vs the electronic format of the questionnaire. However, all patients gave their separate written informed consent before their participation in this ancillary study. At all three visits, the patients filled in the paper format of the WOMAC first, followed by the electronic format.

The validated German paper format of the WOMAC 3.1 with a Likert scale was used[8,9]. The electronic format of the WOMAC was identical to the German questionnaire with an identical Likert scale in a computerized touch screen format, which has been previously shown to have very good agreement with the original paper format in its numeric rating scale format[6]. The QUALITOUCH® data capture method was developed to facilitate patient assessment. The QUALITOUCH® computer program offers a multimedia 3-D interactive interface: the questions are displayed on a 34.3 cm diameter screen as a text and a situational cartoon and are verbalized over the loudspeaker. The questions are answered by touching one of the squares of the Likert scale on the computer screen [Fig. 1(a)]. By using five buttons on screen the patient could exit the questionnaire, get help, have the question repeated or move only one question forward or backward [Fig. 1(b)]. It is therefore possible to leave out one question and move to the next. Furthermore, the help function self activates after 15 s of inactivity and presents the next possible steps to the patient. Patients are not able to see their prior scores.

STATISTICS

Descriptive statistics included the mean of the aggregated scores, the standard deviation and the mean difference between the scores of the paper and computerized

Table I
Patient demographics

	Male ($n = 32$)	Female ($n = 21$)	Total ($n = 53$)
Age in years (mean ± SD)	63.2 ± 10.3	65.7 ± 8.1	64.2 ± 9.5
Height in cm (mean ± SD)	175.5 ± 5.3	162.0 ± 6.8	170.1 ± 8.9
Weight in kg (mean ± SD)	86 ± 15.3	75.7 ± 17.3	81.9 ± 16.8

Table II
Comparative table of the effect of rofecoxib by format (paper vs electronic) and by WOMAC standardized subscale score over time, mean ± SD

Subscale	Format	t_0	t_1	t_2	Effect size t_2 vs t_0
Pain	Paper	39.7 ± 14.5	33.7 ± 15.7	28.6 ± 14.8	0.76
	e-WOMAC	42.3 ± 15.2	34.6 ± 15.6	29.2 ± 14.6	0.88*
Stiffness	Paper	43.4 ± 18.9	36.5 ± 15.7	31.1 ± 20.0	0.63
	e-WOMAC	46.1 ± 22.2	38.6 ± 18.1	33.4 ± 20.8	0.59*
Function	Paper	44.1 ± 14.0	38.4 ± 15.4	32.8 ± 16.2	0.75
	e-WOMAC	43.8 ± 14.3	38.9 ± 16.2	32.0 ± 16.2	0.77*

*Difference between paper vs electronic format scores statistically not significant.

format. To detect possible format (paper vs computer), time (t_0 = baseline vs t_1 = visit 1 at day 7 vs t_2 = visit 3 at day 21), scale (pain vs stiffness, vs function) or gender (male vs female) related effects, a variance analysis by "repeated measures ANOVA" was performed. In addition, the standardized mean difference was used to measure the effect size by WOMAC subscale between t_0 and t_2 (effect size = [{mean score t_2 − mean score t_0}/{pooled SD}] and tested for significance of the paper vs the electronic format. All statistical tests were performed at a significance level of 0.01 or lower to correct for multiple testing. Normality of the distribution was tested by Kolmogorov−Smirnov. All statistical analyses were performed with SAS® StatView® 5.01.

Results

PATIENTS

All consecutive 53 patients recruited in the three participating study centers were included in the analysis and completed both formats of the questionnaire in all three visits. The detailed patient characteristics are shown in Table I. Age and gender were balanced. Forty-three patients (81%) had primary unilateral knee osteoarthritis, 5 (9%) primary bilateral knee OA, 4 (8%) primary idiopathic hip OA and 1 (2%) had secondary hip OA after congenital hip dysplasia. The time needed to answer all questions in the electronic format of the WOMAC 3.1 was 12.9 ± 2.7 min vs 12.5 ± 3.5 min for the paper format (not significant).

WOMAC-SCORES

WOMAC baseline scores by subscale were not significantly different by format (electronic or paper) and the effect size by subscale between t_0 and t_2 was not significantly different by format (Table II). The overall effect size between t_0 and t_2 was 0.71 for the paper version and 0.74 for the electronic version and was statistically not significant. Therefore, the responsiveness by subscale was not significantly different between the electronic and the paper versions of the WOMAC. At visit 1 (t_1 = day 7) pain and stiffness were significantly reduced with rofecoxib, while function was significantly increased ($P < 0.01$, Table III). The mean magnitude of the effect of rofecoxib was continuously increasing over time in all three subscales and in both formats. Comparing baseline (t_0) to visit 2 (day 21), pain decreased by 30%, stiffness by 28% and function increased by 26% irrespective of whether the paper or the electronic format of the WOMAC 3.1 was used (Fig. 2). While the format (paper vs electronic) might have had some influence on the stiffness scale and on the total WOMAC Index, the format × time interaction of WOMAC 3.1 scores between the paper and the electronic formats at t_0 (baseline), t_1 (day 7) and t_2 (day 21) were not significantly different for pain ($P = 0.22$), stiffness ($P = 0.895$), functional ability ($P = 0.542$) and for the total WOMAC Index ($P = 0.508$), indicating that the pattern of changes in WOMAC scores and subscores over time did not differ by format (Fig. 3). The time × format × scale interaction was not significant, indicating that the pattern of changes of the WOMAC scores was the same for both formats and all scales (Table IV). Gender had no influence on the results.

Discussion

This is the first longitudinal study documenting repeated measures with electronic data capturing through patient self-assessment. Electronic data capturing has become increasingly popular for data acquisition in clinical trials. However, the data collected usually refers to laboratory or diagnostic examination values or to patient history and data entry is usually performed by medical or paramedical personnel. In daily clinical care, there is little experience with patient self-assessment using standardized questionnaires and almost no related validated tools exist. One study documents the initial evaluation of an electronic format of the Short Form 36, concluding that electronic data collection is acceptable to patients and feasible in a clinical setting while providing comparable responses to those of the paper format, improving data capture and being immediately available[10].

Table III
Effect of time (t_0, t_1, t_2), format (paper vs electronic) and time × format interaction on the WOMAC subscales and the WOMAC Index

	t_0, t_1*	t_1, t_2*	t_0, t_2*	Format**	Time × format*
Pain subscale	−6.86 ($P = 0.0002$)	−5.27 ($P = 0.0041$)	−12.14 ($P < 0.0001$)	−1.33 ($P = 0.0119$)	$P = 0.22$
Stiffness subscale	−7.73 ($P = 0.0003$)	−4.77 ($P = 0.0243$)	−12.5 ($P < 0.0001$)	−2.73 ($P = 0.0078$)	$P = 0.895$
Function subscale	5.32 ($P = 0.0018$)	6.25 ($P = 0.0003$)	11.58 ($P < 0.0001$)	0.26 ($P = 0.5770$)	$P = 0.542$
WOMAC 3.1 Index	−6.64 ($P < 0.0001$)	−5.43 ($P = 0.0008$)	−12.07 ($P < 0.0001$)	−1.27 ($P = 0.0049$)	$P = 0.508$

*Significant if $P < 0.0033$. **Significant if $P < 0.01$.

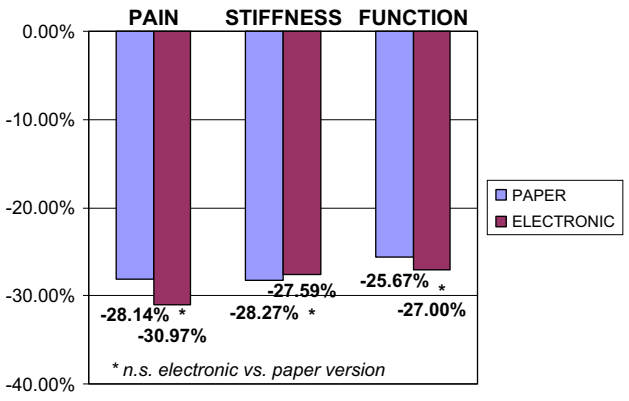


Fig. 2. Evolution of WOMAC 3.1 subscales with rofecoxib one tablet once daily over 3 weeks, paper vs electronic evaluation.

Patient self-assessment by electronic data capturing presents numerous advantages in clinical care: the data collection is standardized and the impact of potential external influences, which may vary in nature from visit to visit, is limited; the data may be collected anonymously across departments, hospitals and medical practices allowing for constant optimization of patient management techniques by detecting outliers regarding treatment success. When performed in the waiting-room, self-assessment makes the best use of the patient's and the physician's time and is a valuable contribution to the patient–physician interaction, especially in the decision-making process of treatment adaptations. In contrast with paper questionnaires which are archived in the patient's file, electronic data allow for easy treatment effect monitoring at a glance. With the e-WOMAC, the patient's progress is documented for the three subscales: pain, stiffness and function. Multiple assessments over time are displayed on one page displaying the change over time in an easy to read graph. The e-WOMAC data collection by patient self-assessment with a QUALITOUCH® touch screen interface has been shown to have very good agreement in all three subscales with the original paper format of the questionnaire[6]. In another study, the patient preference for the electronic vs the paper format of the questionnaire was documented: although 54% of the patients had no experience with computers at all, only 9% preferred the paper format, 91% either preferring the computer format (51%) or being indifferent (38%). Ninety-four percent of the patients declared that the 3-D environment presented (text, sound and cartoon) was helpful[7].

This study shows that e-WOMAC is responsive to treatment over time with regard to pain, stiffness and function. In addition, no significant difference was found while comparing the degree of responsiveness by subscale between the electronic and the paper format of the

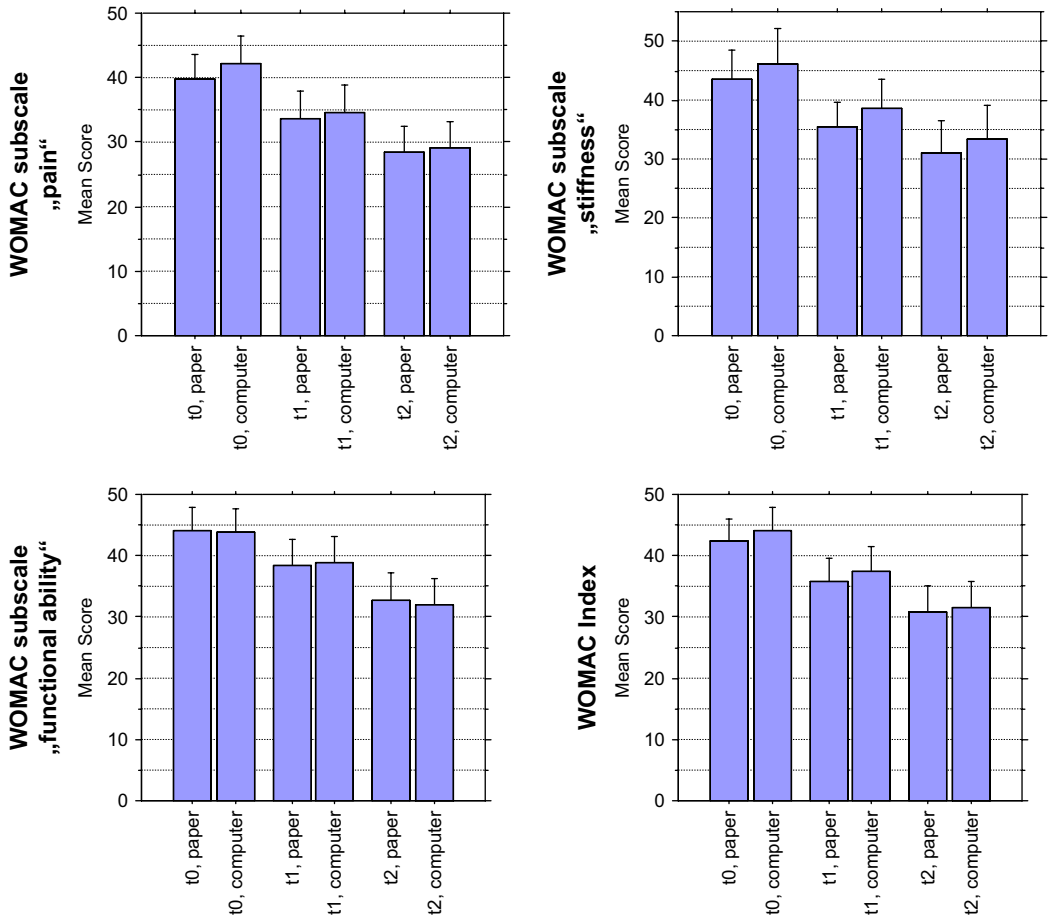


Fig. 3. Changes in mean WOMAC subscales and in WOMAC Index (±95% confidence intervals), by visit at t_0 (baseline), t_1 (day 7) and t_2 (day 21), paper vs computer format, with rofecoxib one tablet once daily.

Table IV
WOMAC Index: levels of significance with repeated measures ANOVA with one, two or three factors

Effect	*P* value	Interpretation
Time* (t_0, t_1, t_2)	<0.0001	Rofecoxib significantly improved WOMAC score over time
Format* (paper vs e-WOMAC)	0.0049	Format had an influence on WOMAC score
Scale* (pain vs stiffness vs function)	0.0183	Subscale had no influence on WOMAC score
Time × format*	0.5077	Changes of WOMAC scores over time were the same for both formats
Time × scale*	0.7229	Changes of WOMAC scores over time were the same for all WOMAC subscales
Time × format × scale*	0.7703	Patterns of change in WOMAC scores over time did not differ by version and scale

*Significant if $P < 0.01$.

questionnaire. This suggests that the e-WOMAC is as responsive as the original paper format. As the main endpoints of the SVIS study were to document the effects of rofecoxib on Quality of Life (measured by the SF-12) and disease specific symptoms (measured by the WOMAC paper questionnaire) and because at the time of the initiation of the SVIS study the e-WOMAC was not completely validated, we renounced to randomize for the two formats (paper and electronic) and asked all patients to fill in the paper format of the questionnaire consistently before they filled in the electronic format, accepting thereby a systematic error in the validation procedure. As a single parameter, the format seemed to have an influence on the WOMAC score, the significance being driven by the two questions related to pain (Table V). In this study, the overall effect size between t_0 and t_2 reached 0.71 when measured with the paper version of the WOMAC and 0.74 with the electronic version. An effect size between 0.2 and 0.5 is considered as small but clinically meaningful, while a large effect size is estimated at being 1.0 or more[11]. Therefore the observed effect size of rofecoxib between t_0 and t_2 should be considered not only statistically significant but also clinically relevant. In contrast the difference in effect size of 0.03 observed between the paper and the electronic version of the WOMAC is statistically non-significant and should be considered as clinically irrelevant. This holds true for all three subscales of the WOMAC, the largest observed difference in effect size between formats being 0.12. Furthermore, the paper and the electronic formats of the WOMAC have proven to be very similar for the monitoring of treatment effects and under the premises that the choice for the paper or the electronic format is made upfront and carried out throughout the timecourse of the observation, both formats can be considered as equivalent. In the meantime, another study with correct block randomization has demonstrated the patient's preference for the electronic format and the perfect interchangeability of the paper and the electronic formats[7].

Table V
Influence of the version in relation with the section of the questionnaire

Questions addressing	Number of questions	Significance level
Pain	5	0.012
Stiffness	2	0.0078
Function	17	0.577

*Significant if $P < 0.01$.

In patients with symptomatic OA at the hip or knee treated with rofecoxib, the paper and the electronic format of the WOMAC 3.1 showed similar effect sizes and were equally suitable for the longitudinal monitoring of the effects of drug treatment and the detection of clinically meaningful changes. The future successful use of the e-WOMAC by the QUALITOUCH® method in medical care will depend on its integration in the daily processes of patient management at the primary care physician level.

References

1. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol 1986;5(2):231–41.
2. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833–40.
3. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al*. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24(4):799–802.
4. Pincus T, Sokka T. Quantitative measures for assessing rheumatoid arthritis in clinical trials and clinical care. Best Pract Res Clin Rheumatol 2003;17(5):753–81.
5. Theiler R, Bischoff HA, Good M, Uebelhart D. Rofecoxib improves quality of life in patients with hip or knee osteoarthritis. Swiss Med Wkly 2002;132(39–40):566–73.
6. Theiler R, Spielberger J, Bischoff HA, Bellamy N, Huber J, Kroesen S. Clinical evaluation of the WOMAC 3.0 OA Index in numeric rating scale format using a computerized touch screen version. Osteoarthritis Cartilage 2002;10(6):479–81.
7. Bischoff-Ferrari HA, VonDechend M, Bellamy N, Theiler R. Validation and patient acceptance of a computer touch screen version of the WOMAC 3.1 Osteoarthritis Index. Ann Rheum Dis July 1 2004; PMID:15231508.
8. Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, *et al*. [Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) Arthrosis Index]. Z Rheumatol 1996;55(1):40–9.
9. Bellamy N. Personal communication of the German version of the WOMAC 3.1 with Likert scale.
10. Wilson AS, Kitas GD, Carruthers DM, Reay C, Skan J, Harris S, *et al*. Computerized information-gathering in specialist rheumatology clinics: an initial evaluation of an electronic version of the Short Form 36. Rheumatology (Oxford) 2002;41(3):268–73.
11. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, NJ: L. Erlbaum Associates; 1988: p. xxi, 567.